Quadratic Regression

Oct. 25 2019

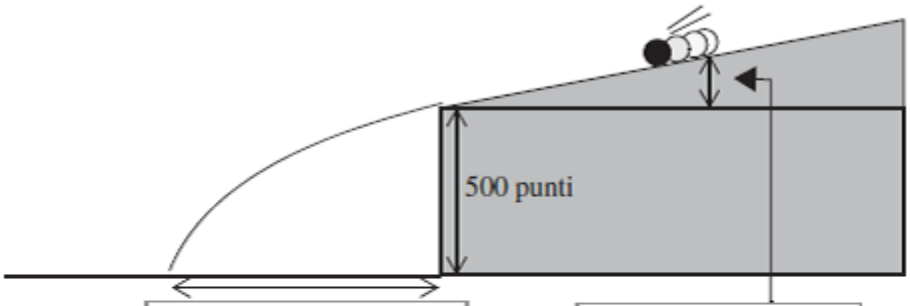
Summary

- Sometimes there is actually a quadratic relationship between two variables: $\mu(Y|X) = \beta_0 + \beta_1 X + \beta_2 X^2$
- Fit with `lm(response ~ poly(explanatory, degree = 2, raw = TRUE), data = data)`

Galileo's data on falling things (Sleuth3 Case study 10.1.1)

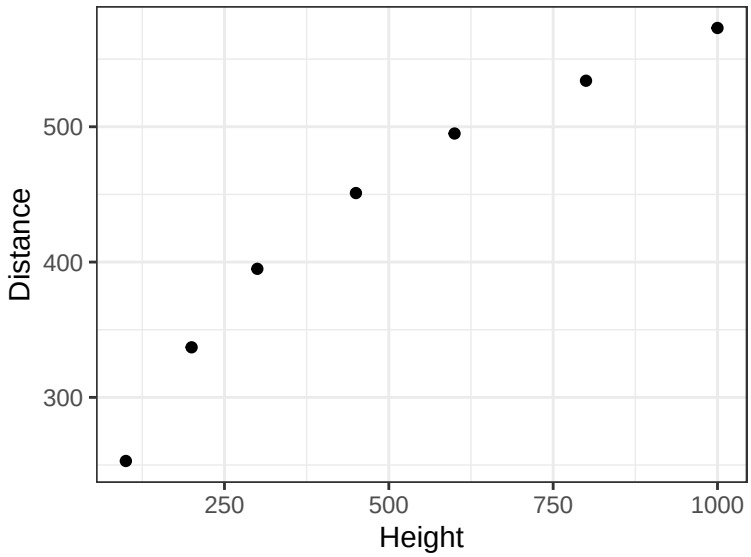
- Galileo showed that the trajectory of a body falling with horizontal velocity is a parabola.
- Rolled ink-covered bronze ball down an inclined plane
- Measured vertical height and horizontal distance in units of punti (1 punto is 169/180 mm)

DISPLAY 10.1 Galileo's experimental results



Horizontal Distance (punti)	Initial Height (punti)
253	100
337	200
395	300
451	450
495	600
534	800
573	1,000

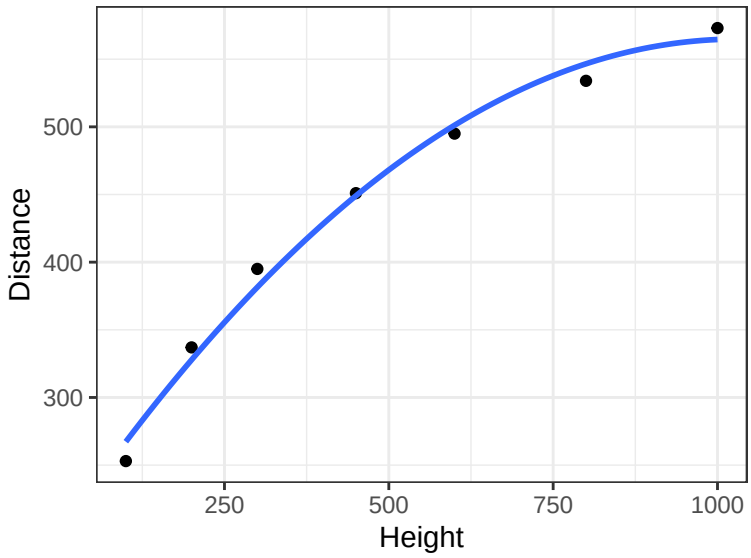
```
ggplot(data = galileo, mapping = aes(x = Height, y = Distance)) +
  geom_point() +
  theme_bw()
```



```r
quadratic_fit <- lm(Distance ~ poly(Height, degree = 2, raw = TRUE), data = galileo)
summary(quadratic_fit)
```

```
##
## Call:
## lm(formula = Distance ~ poly(Height, degree = 2, raw = TRUE),
##     data = galileo)
##
## Residuals:
##       1       2       3       4       5       6       7
## -14.308   9.170  13.523   1.940  -6.177 -12.607   8.458
##
## Coefficients:
##                                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)                         1.999e+02  1.676e+01  11.928 0.000283 ***
## poly(Height, degree = 2, raw = TRUE)1  7.083e-01  7.482e-02   9.467 0.000695 ***
## poly(Height, degree = 2, raw = TRUE)2 -3.437e-04  6.678e-05  -5.147 0.006760 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.64 on 4 degrees of freedom
## Multiple R-squared:  0.9903, Adjusted R-squared:  0.9855
## F-statistic:   205 on 2 and 4 DF,  p-value: 9.333e-05
```

```r
ggplot(data = galileo, mapping = aes(x = Height, y = Distance)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y ~ poly(x, degree = 2, raw = TRUE), se = FALSE) +
  theme_bw()
```



1. What model did we fit?

2. What is the estimated equation describing the relationship between height and "population" mean distance?